Are all kids alike? The magnitude of individual differences in personality characteristics tends to increase from early childhood to early adolescence

Accepted in European Journal of Personality

Are all kids alike? The magnitude of individual differences in personality characteristics tends to increase from early childhood to early adolescence

René Mõttus
University of Edinburgh, UK
University of Tartu, Estonia

Christopher J. Soto
Colby College, Maine, US

Helena R. Slobodskaya
Institute of Physiology and Basic Medicine, Novosibirsk, Russia
Novosibirsk State University, Russia

Authors' Note:
This research was supported in part by a faculty research grant from Colby College to Christopher J. Soto and by grants from the Russian Scientific Foundation #16-18-00003 and Russian Foundation for Basic Research #16-06-00022 to Helena R. Slobodskaya.

Abstract

Do individual differences in personality traits become more or less pronounced over childhood and adolescence? The present research examined age differences in the variance of a range of personality traits, using parent-reports of two large samples of children from predominantly the United States and Russia, respectively. Results indicate (a) that individual differences in most traits tend to increase with age from early childhood into early adolescence and then plateau, (b) that this general pattern of greater personality variance at older childhood age is consistent across the two countries, and (c) that this pattern is not an artifact of age differences in means or floor/ceiling effects. These findings are consistent with several (noncontradictory) developmental mechanisms, including youths' expanding behavioral capacities and person-environment transactions (corresponsive principle). However, these mechanisms may predominantly characterize periods before adolescence, or they may be offset by countervailing processes, such as socialization pressure towards a mature personality profile, in late adolescence and adulthood. Finally, the findings also suggest that interpreting age-trajectories in mean trait scores as pertaining to age differences in a typical person may sometimes be misleading. Investigating variance should become an integral part of studying personality development.

Keywords: little six; variance; development; corresponsive principle; transactions

Are all kids alike? The magnitude of individual differences in personality characteristics tends to increase from early childhood to early adolescence

How do personality traits differ across the life span? One way to address this fundamental question is by examining age differences in the *mean levels* of personality traits, which are generally interpreted as representing the average age trend, or the age trend pertaining to a typical person. Another approach is to examine age differences in the *variance* of traits, which indicate whether individual differences in personality become more or less pronounced (or neither) across the life span. Previous research has mostly examined age differences in personality trait mean levels (e.g., Donnellan & Lucas, 2008; Roberts, Walton, & Viechtbauer, 2006; Soto, John, Gosling, & Potter, 2011), whereas the possibility of age differences in personality variance has received comparatively little attention (Mõttus et al., 2015; Small, Hertzog, Hultsch, & Dixon, 2003). In order to address this discrepancy, the present research examines whether the variability of personality traits differs systematically by age from early childhood through late adolescence. We did this by analyzing parent-reports describing personality traits in two large, cross-sectional samples of youths: one predominantly from the United States and one from Russia.

Possible Mechanisms for Age Differences in Variance

Do individual differences in personality traits become more or less pronounced across childhood and adolescence? There are several plausible mechanisms by which they might become more pronounced. For example, youths develop new cognitive (Flavell, Miller, & Miller, 1993; Inhelder & Piaget, 1958), emotional (Murphy, Eisenberg, Fabes, Shepard, & Guthrie, 1999), and self-regulatory (Demetriou, 2000; Gestsdottir & Lerner, 2008) capacities throughout childhood and adolescence. This gradual expansion of the behavioral repertoire may provide older youths with more ways to express their distinctive personality, and may also provide parents, peers, and other observers with more cues to detect individual differences between youths. For example, gains in hypothetical thinking and abstract reasoning capacities from childhood into adolescence, as the result of ongoing cognitive maturation and academic instruction (Flavell, Miller, & Miller, 1993; Inhelder & Piaget, 1958), may allow older youths to more clearly express individual differences in Openness to Experience. An increase in the number of options available for expressing and detecting personality traits could then contribute to greater observed personality variance at older ages.

A second possible mechanism concerns identity development. Over time, especially during adolescence, youths typically develop clearer and better differentiated self-concepts (Byrne & Shavelson, 1996; Donahue, 1994; Erikson, 1968; Harter, 1999, 2006; Harter & Monsour, 1992; Marsh, 1989; Marsh & Ayotte, 2003; Montemayor & Eisen, 1977; Soto, John, Gosling, & Potter, 2008) and narrative identity (McAdams & McLean, 2013). They also gradually become more psychologically autonomous from their parents (Greenberger, 1982, 1984; Steinberg & Silverberg, 1986). If older youths have a clearer sense of their distinctive identity, as well as a greater capacity to act in ways consistent with their identity, this may further promote the expression of individual differences in personality traits.

At the same time that youths may experience these intrapersonal changes in their behavioral capacities and identity, they might also experience changing relationships with their social environment. For example, as youths spend less time closely supervised by parents and teachers (Patterson & Stouthamer-Loeber, 1984), they gain greater freedom to selectively seek out or create situations that allow them to express their personality traits (situational selection), and to avoid situations that are inconsistent with their personality (situational attrition). Over time, youths' self-selected situations may deepen the personality characteristics that led to them in the first place, a phenomenon referred to as the corresponsive principle of personality development (Caspi et al., 2005; Roberts et al., 2008). For example, extraverted children may

seek out socially stimulating environments, which may further enhance their social skills and thereby accentuate their Extraversion even more over time. In contrast, introverted children may avoid socially stimulating environments and thereby deprive themselves from practicing relevant skills, which may further lessen their social ambitions and thereby decrease their level of Extraversion. Over time, these person-environment transactions would tend to deepen individual differences in personality traits, thereby increasing personality variance.

In contrast with the developmental processes, reviewed above, that might lead youth personality variance to increase with age, at least one notable mechanisms—socialization—may lead to *less* pronounced individual differences in personality traits at older ages. Most youths are encouraged to behave in certain socially acceptable ways, and discouraged from behaving in other, unacceptable ways. For example, children in many cultures are encouraged to help, respect, and share with others, and to complete assigned tasks and fulfill their obligations to others; they are generally discouraged from harming others, and from tantrums and other excessive displays of negative emotion (Colby et al., 1983; Eisenberg, Carlo, Murphy, & van Court, 1995; Erikson, 1968; Eisenberg & Morris, 2004). In personality trait terms, children are encouraged to become psychosocially mature by developing high Agreeableness and Conscientiousness, as well as low Neuroticism (Caspi et al., 2005; Digman, 1997). Some youths begin life temperamentally predisposed toward this mature personality profile, while others have a greater psychological distance to travel to reach there (Clark & Watson, 2008; Krueger & Johnson, 2008; Rothbart, Ahadi, & Evans, 2000). To the extent that socialization processes make youths more alike over time—by pulling them toward the trait profile of a typical, mature person—they will tend to mute temperamental differences and thus gradually decrease personality variance.

Previous Research Examining Age Differences in Personality Traits

As noted above, much previous research has examined mean-level age differences in personality traits across the life span. Collectively, this evidence indicates that most individuals become more agreeable, conscientious, and emotionally stable across adulthood (e.g., Donnellan & Lucas, 2008; Roberts et al., 2006), although the pattern is more varied at the level of more specific traits (Mõttus et al., 2015). During childhood and adolescence, however, the general trends toward greater psychosocial maturity may be disrupted by temporary dips in Agreeableness, Conscientiousness, and Openness to Experience, as well as increases in Neuroticism (among girls only), during the years surrounding puberty (Denissen, van Aken, Penke, & Wood, 2013; Soto, 2016; Soto et al., 2011; van den Akker, Dekovic, Asscher, & Prinzie, 2014).

Much less research has investigated age differences in personality variance, either in adulthood or in childhood and adolescence. Although some studies on personality development in youths report age-specific standard deviations as descriptive statistics, few have directly tested for age differences in personality variance. For example, among the 14 studies of mean-level youth personality development reviewed by Denissen et al. (2013), only one explicitly addressed age differences in variability, although several reported trait score standard deviations for different age groups among other statistics. Specifically, Allik et al., (2004), studying Estonian children from age 12 onwards, noted that there was "no systematic trend in standard deviations: older children were not more variable in their personality traits than younger children" (p. 450). A similar conclusion could be drawn by inspecting the standard deviations reported by De Haan et al. (2013) for Belgian children aged between 9 and 12 years. In contrast, inspection of the descriptive statistics reported by two other studies conducted with Dutch (Aa et al., 2009) and American (Johnson et al., 2007) adolescents and emerging adults suggests a tendency—not discussed by the authors—toward slightly decreasing personality variance with age.

In adulthood, one study that specifically focused on age differences in the magnitude of individual differences (Mõttus et al., 2016) examined variance of the Five-Factor Model of

personality (FFM; McCrae & John, 1992) domains and their 30 facets using self-report and informant-report data from Estonia, the Czech Republic, and Russia. This study did not find systematic age differences in variance for any domain or facet, either when comparing late adolescents (ages 16 to 20) to emerging adults (ages 21 to 25) or when comparing young adults (ages 20 to 30) to middle-aged adults (ages 50 to 60).

To our knowledge, the present research is the first to conduct a similarly systematic examination of age differences in personality variance from early childhood through late adolescence. This research should be considered exploratory, in that it is unclear whether we should expect the amount of personality variance to remain consistent across these developmental periods, as Allik et al. (2004) and Mõttus et al. (2016) observed from adolescence onwards, or whether the developmental trends in behavioral capacities, identity, person-environment transactions, and socialization reviewed above might lead to age-related increases or decreases in personality variance during the first two decades of life.

Overview of the Present Research

In order to better understand trajectories of the magnitude of individual differences in personality characteristics, the present research examined age differences in trait variance at three levels of abstraction: broad personality domains, narrower facets, and specific nuances (McCrae, 2015). Specifically, Study 1 investigated cross-sectional age differences in the variance of the Little Six (Soto & John, 2014) personality domains—Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Activity—from early childhood through adolescence in a sample of children from predominantly English-speaking nations (mostly the United States). Recent theoretical and empirical work suggests that these six broad domains may particularly well capture the basic structure of youth personality traits during childhood and adolescence (De Pauw, in press; De Pauw & Mervielde, 2010; De Pauw, Mervielde, & van Leeuwen, 2009; Shiner & DeYoung, 2013; Soto, 2016; Soto & John, 2014). Study 2 attempted to replicate the key findings of Study 1 in a different culture (the Asian part of Russia), using a questionnaire of child personality that measures 15 facet-level personality traits more specific than the broad Little Six. In both studies, we also examined personality variance at the level of single questionnaire items, which represent specific personality "nuances" that provide incremental information beyond broad trait domains and mid-level facets (McCrae, 2015; Mõttus, Kandler, Bleidorn, Riemann, & McCrae, 2017; Soto, 2016). For example, Mõttus and colleagues (2017) demonstrated that the unique variance of item-level nuance traits is often heritable, stable over time, agreed upon by different raters, and predictive of external criteria. We did not have any *a priori* reason to expect the findings to differ across the two studies, also representing children from different cultural backgrounds.

The present research has three key strengths that make it particularly well-suited for examining the development of personality variance. First, both studies were based on large samples of youths spanning from early childhood through adolescence (total $N = 17{,}996$), which allowed us to examine age differences in personality variance with considerable precision. Large and similarly (often identically) sized age groups mitigated possible problems arising from comparing small/unequally sized groups. Second, both studies examined parent-reports of children's personality characteristics; compared with youth self-reports, our use of parent-reports helped us separate age differences in personality variance from age differences in youths' capacity to provide meaningful personality reports (Soto et al., 2008). Third, differences between the two studies' samples and measures allowed us to test whether the key findings generalize across cultural contexts and levels of trait abstraction. Also, the analyses were based on a wide range of constructs, helping us to mitigate problems specific to particular measurements (e.g., ceiling/floor effects of some question or scale scores).

Study 1

Method

Sample and Procedure

Participants were mostly the parents or guardians of 16,000 children, adolescents, and young adults between the ages of 3 and 20 years old (for 9% of raters their relationship with the child was not known). The sample of target children included 500 males and 500 females in each of 16 age groups: each individual year of age from 3 to 17, plus a combined 18-20 year old group. Initially, the sample was larger (N = 24,373); the 1,000 children in each age band were randomly selected from this initial sample, with the purpose of having equally sized and sex-balanced groups for each age. In terms of ethnicity, 78% of the target children were described as White/Caucasian, 4% as Black/African American, 4% as Hispanic/Latino, 3% as Asian/Asian American, 1% as Native American/American Indian, 2% as another ethnicity, and 8% as multiple ethnicities. Approximately 83% resided in the United States, 7% in the United Kingdom or Ireland, 6% in Canada, and 4% in Australia or New Zealand. Most of the raters (89%) were mothers.

Participants were visitors to a non-commercial website, personalitylab.org, who volunteered to anonymously complete a version of the common-language California Child Q-set (CCQ; Caspi et al., 1992) in exchange for automatically generated feedback about their child's personality. Additional details about the present sample and procedure are available from two previous studies that examined the multidimensional structure of the CCQ (Soto & John, 2014) and mean-level age differences in the Little Six (Soto, 2016); however, neither of these previous studies examined age differences in the variability of youth personality traits.

Measure

The common-language CCQ (Caspi et al., 1992) was developed to allow researchers, clinicians, parents, and other observers to comprehensively rate youths' personal characteristics. Its items are 100 statements that respondents use to evaluate the target child. The common-language CCQ was modified in two ways for the present research (see Soto & John, 2014). First, participants independently rated each item on a scale ranging from 1 (*extremely uncharacteristic*) to 9 (*extremely characteristic*) rather than sorting the items into a fixed Q-sort distribution (see Block & Block, 1980). Second, specific words or phrases (e.g., "child," "kid") in 12 items were replaced (e.g., with "person") so that these items could be applied to adolescents and young adults as well as children.

Soto and John (2014) distinguished 94 CCQ items that primarily assess a personality characteristic from 6 items that primarily assess an overt physical characteristic (e.g., physical appearance) or social effect (i.e., another person's behavior toward the target child). Soto (2016) then developed scales that use 67 of the personality-focused CCQ items to assess the Little Six youth trait domains. Following this previous work, the present research examines scale-level analyses of the CCQ-Little Six scales, as well as item-level analyses of the 94 personality-focused CCQ items. In the present sample, the alpha reliability coefficients of the CCQ-Little Six scales were .83 for Extraversion (9 items), .93 for Agreeableness (22 items), .86 for Conscientiousness (10 items), .87 for Neuroticism (15 items), .63 for Openness (6 items), and . 73 for Activity (5 items; see Soto, 2016).

Data analysis

Quantitative analyses were carried out using R statistical language (R Core Team, 2017). The scripts and data are made publicly available in Online Supplementary Material (osf.io/aeu2b).

Results and Discussion

Initial Distribution of Scores

Before the main analyses, it was important to investigate whether the distributions of the Little Six domain scores were systematically more truncated in some age groups than in others, because this could produce confounded results. This is because instrument-specific floor and ceiling effects inevitably reduce the observable variance, whereas the true population variance may not be truncated. Indeed, the distributions of scale scores appeared skewed, especially for Extraversion, Activity, and Openness (reflecting ceiling effects); distributions of all six domains for all age groups are given in the Online Supplementary Material (osf.io/aeu2b). Generally, the skewness was more pronounced at earlier ages. We therefore created parallel reduced versions of the scales by omitting all items with absolute skewness values higher than 0.5 at age 3; for Activity and Openness we had to adopt more lenient criteria of 0.85 and 1.10, respectively, in order to retain at least three items in both scales (so that latent trait models could be identified). The reduced Neuroticism, Extraversion, Agreeableness, and Conscientiousness scales included nine, three, nine and six items, respectively. The distributions of reduced scale scores in all age groups are also given in the Online Supplementary Material, alongside the distributions of full scales. Generally the skewness was smaller in the reduced scales, although it tended to increase for Openness in older age groups. We will present the results for both full scales (based on all items selected for the respective scales) and reduced scales (excluding the most skewed items), in order to test whether the findings were robust with regard to skewness.

Measurement Invariance (MI)

Before comparing age groups in the variance of trait scores, it was also important to ascertain that the trait scales measured the same constructs at different ages—that is, that the measurements were invariant. For a thorough treatment of MI and its implications for studying age differences in personality traits see Mõttus and colleagues (2015).

For each Little Six scale (full and reduced), we created a unidimensional confirmatory factor analysis (CFA) model, whereby its intended items defined a single reflective latent trait (i.e., a trait that is postulated as the common cause of its defining items; Edwards & Bagozzi, 2000). As is common in such cases (e.g., Mõttus et al., 2015), there was a pervasive lack of local independence, with items showing residual correlations after being conditioned on the latent trait. We allowed as many item residuals to correlate as necessary to achieve a confirmatory fit index (CFI) value of at least .95 and a root mean square error of approximation of no higher than 0.08 (van de Schoot, Lugtig, & Hox, 2012). The residual correlations were identified iteratively until sufficiently good fit was obtained: in each iteration, the item pair with the highest modification index for its residual correlation was allowed to covary. In many cases, we had to allow for numerous residual correlations to achieve acceptable model fit (e.g., 56 for the full Agreeableness scale). Fit indices for models without and with residual correlations are given in the Online Supplementary Material (for each model, the number of residual correlations can be worked out by comparing the degrees of freedom). For reference, the main findings based on models *without* allowing for residual correlations are also reported in Online Supplementary Material (Figure 1SA).

Next, we implemented these models as multi-group models by using the target child's age as the grouping variable (ages 18 to 20 were combined into a single group). We carried out standard MI tests by first allowing all model parameters to vary across the groups, but then sequentially constraining all (1) factor loadings, (2) residual correlations and (3) residual variances to be group-invariant (Meredith, 1993). Establishing this degree of MI would allow us to compare factor variances across groups. Although we did not plan to compare factor means, for the sake of completeness we also tested for intercept equality across groups. By comparing each model to the previous, less constrained model we could test for the equality of the

constrained parameters; the criterion for a lack of MI was a decrease in CFI value of more than .01 (Cheung & Rensvold, 2002).

All but one scale met the criteria of weak MI (equality of factor loadings); for the reduced Extraversion scale, CFI dropped by .05 after constraining factor loadings equal across the age groups. This was caused by the item "Likes to be by him/herself; enjoys doing things alone". We set the loading of this item free across age groups, thereby only establishing/assuming partial MI for this scale. Constraining residual correlations equal did not entail a notable decrease in model fit for any scale. Equality of residual variances across groups held for both the full and reduced scales of Neuroticism, Agreeableness and Conscientiousness. The decrease in CFI was substantial for the full Activity scale (.10), but ranged between .03 and .05 for the other scales.

Because the lack of MI at the level of residual variances was not substantial in most cases, we used the models with loadings, residual correlations and residual variances constrained equal for comparing latent factor variances across age groups (with the exception of the reduced Extraversion scale, for which parameters of one item varied freely across the groups). However, we also performed age group comparisons in latent trait variance using models with only factor loadings and residual covariances constrained equal, and obtained results similar to those reported below (see Figure 1SB in Online Supplementary Material). This suggests that the violations of MI did not affect our primary conclusions and the findings were robust across modeling approaches. This conclusion was further supported by the similar results of item-level analyses that will be described below: none of these measurement model-related concerns is relevant for item-level analyses.

We note that none of the scales met the criterion for strong MI (intercept equality), with decreases in CFI ranging from .02 to .16 (median .06). This would have complicated mean-level comparisons, suggesting that age differences were often specific to individual items (nuances) as has been found previously (Mõttus et al., 2015). However, mean-level comparisons were not focal for our study.

Age Differences in Little Six Domain Variance

We analyzed age differences in trait variance by comparing squared variance estimates from multi-group CFA models. This straightforward approach allowed each age level to have a standard deviation and its confidence intervals for each age group.

Analyses of latent trait factors. Standard deviations of the latent Little Six trait domains, along with their 95% confidence intervals, are presented in Figure 1; the standard deviations of the observed sum-scores of the Little Six scales for all age groups are given in the Online Supplementary Material (Figure 1SC). For all domains except reduced Extraversion, standard deviations tended to increase with age across childhood (from approximately ages 3 to 11, 12 or 13); for Neuroticism (especially the reduced scale-based factor scores), a slight increase continued throughout adolescence. For both the full and reduced versions of Activity, Openness, Agreeableness, and Conscientiousness, the confidence intervals of the standard deviation at age 3 did not overlap with those from ages 6 or 7 onwards. This pattern also held for the full versions of Neuroticism and Extraversion, but in the reduced versions it only applied for Neuroticism after age 11 and never for Extraversion. Generally, the trend for increasing variance with age was stronger in full scales than in reduced scales, suggesting that the skewness of items at earlier ages inflated—but almost certainly did not fully explain—the pattern. The pattern of increasing variance appeared to be consistently most pronounced for Activity and Conscientiousness.

For reduced scales, we also repeated the analyses by incorporating Little Six scales in a single CFA model instead of specifying six separate models. The pattern of results was similar to the findings reported above, with variances of latent traits generally increasing with age except for Extraversion (see Online Supplementary Material). Of note is that constraining the covariance of the latent traits equal across age groups did not deteriorate model fit, suggesting that the inter-correlations of the Little Six traits could be considered invariant at different ages.

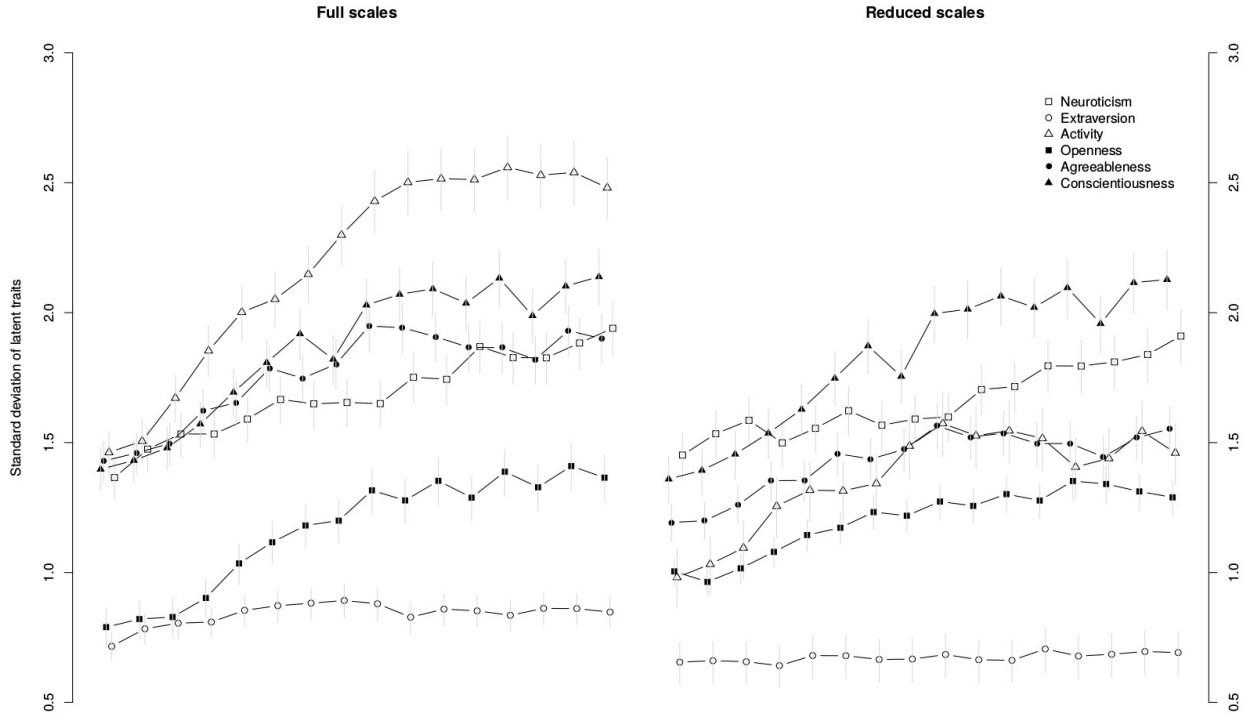

Figure 1. Standard deviations of the Little Six latent factor scores (vertical lines indicate 95% confidence intervals). "Full scales" indicate scores of scales with all items included, whereas "Reduced scales" omitted items with the most skewed distributions.

Age Differences in Little Six Distribution Shapes and Central Tendencies

Figures 2 and 3, respectively, present the kernel density plots describing the distributions of full and reduced (observed, not latent) scale scores for each Little Six domain at four ages: 3, 8, 13, and 18 (distributions for all other ages are given in the Online Supplementary Material). Consistent with Figure 1, these density plots show that trait distributions generally became wider and lower in older ages, especially up to age 13, indicating greater trait variability with age from early childhood into early adolescence.

These plots also depict the age differences in central tendencies and suggest that the differences in variance were not artifacts due to shifts in the locations of the distributions (i.e., shifts away from floor or ceiling effects that reduced variance at younger ages). For example, the small U-shaped mean-level age trends for Openness, Agreeableness, and Conscientiousness (Soto, 2016) did not reflect uniform downward-then-upward shifts in the distributions of these trait domains; instead, both very low and very high levels of these domains became more frequent at older ages. Moreover, although Extraversion and Activity showed very similar patterns of pronounced mean-level decreases with age (Soto, 2016), these two domains showed very different age trends in variance. Specifically, Extraversion showed a relatively consistent amount of variance across all ages, whereas Activity showed substantially increasing variance, reflecting a greater frequency of very low scores (and possibly *very* high scores, in the case of the reduced Activity scale) at older ages. Thus, the overall pattern of greater variability in the Little Six domains at older ages could not be explained by shifts away from floor or ceiling effects. Neither could the pattern be explained by increasing mean scores: the means of most domains decreased, whereas variances increased. In contrast, it appears that age differences in central tendencies could sometimes (clearly in case of Activity) be partly accounted for by age differences in variances (e.g., increases in the number of low scores) rather than the other way around.

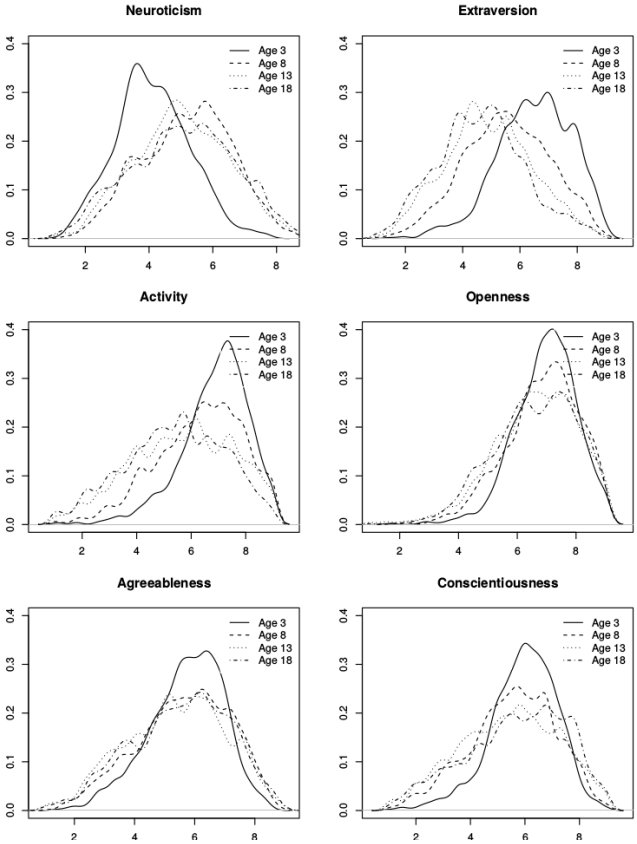

Figure 2. Density distributions of the full scale (all items included) scores (horizontal axis) for four ages. Other ages were omitted for ease of visual inspection. Vertical axis represents density.

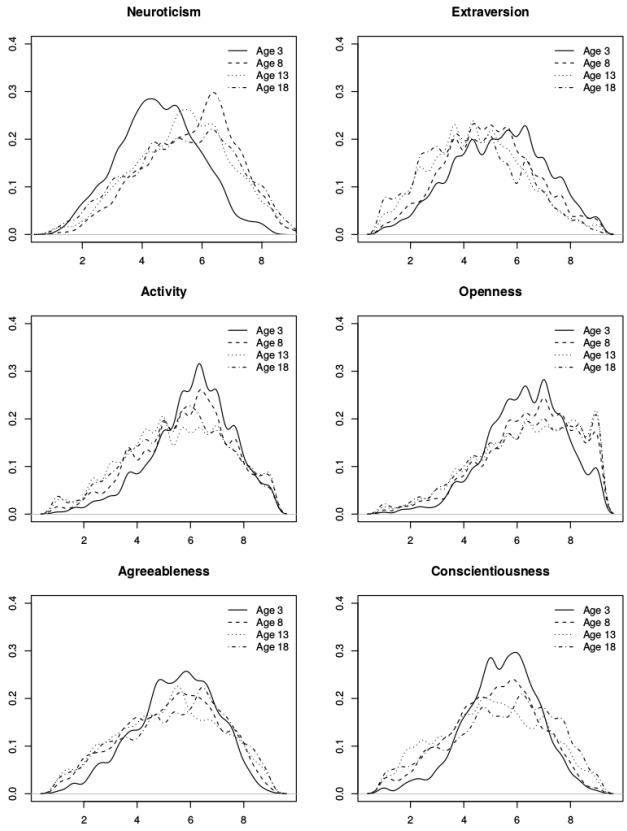

Figure 3. Density distributions of the reduced scale (items with most skewed distributions omitted) scores (horizontal axis) for four ages. Other ages were omitted for ease of visual inspection. Vertical axis represents density.

Age Differences in Item-Level Nuance Trait Variance

The analyses of latent and observed domain-level trait variances suggested that individual differences in the Little Six tend to become more pronounced with age across childhood and into early adolescence. Next, we tested the consistency of this pattern across specific personality nuances, as represented by single CCQ items. First, for the 67 items included in the Little Six scales, we calculated standard deviations for each age group and then linked these with age using Spearman's rank-order correlation, with positive correlations indicating tendencies for increasing variance with age. The median correlation was 0.93 (interquartile range = 0.75 to 0.96); 61 (91%) of the correlations were above 0.50, and only two were negative (both belonging to Extraversion: “Shows their thoughts and feelings, but doesn’t talk much about them” and “Likes to keep their thoughts and feelings to themself”).

For the 27 items not included in the Little Six scales, the same tendency appeared (median correlation = .79, interquartile range = .63 to .95): 21 (78%) correlations were higher than 0.5, and only three were negative ("Has specific habits or patterns of behavior", "Is careful not to get hurt (physically)", and "Has emotions that don't seem to fit the situations"). The averages of the standard deviations of these 27 items are shown in Figure 4; this figure reveals a tendency similar to that observed for the Little Six traits, whereby variance generally increased until the early teenage years and then plateaued.

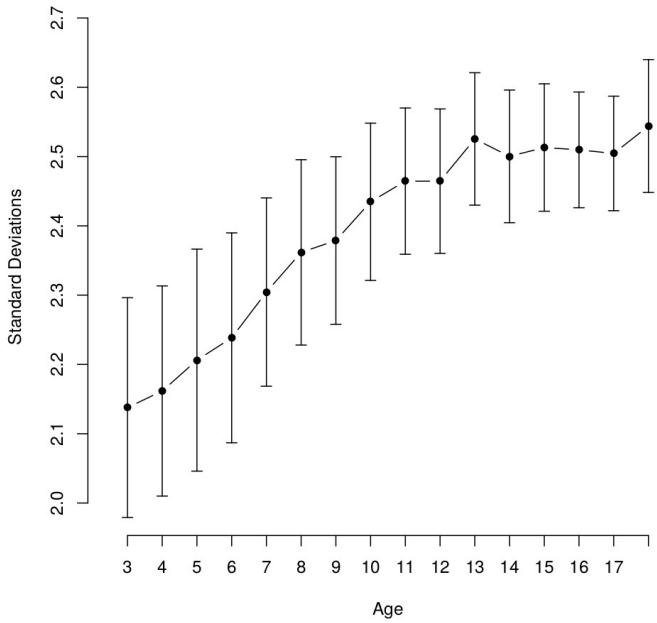

Figure 4. Average standard deviations for the 27 items not included in the Little Six Scales (vertical lines indicate 95% confidence intervals).

The pattern of increasing variance was thus consistent across item-level nuance traits, with only a few exceptions. This attests the robustness of the findings across possible ways of combining items into traits and arranging measurement model parameters. However, as was observed at the level of the Little Six domains, age differences in variance were somewhat inflated by the distributions of many items being skewed at earlier ages. For example, there was a (Spearman) correlation of 0.38 ($p < .001$) between the skewness of the 94 items at age 3 and the correlation of their standard deviations with age. (The correlation was 0.36, $p < .001$, when the absolute difference of an item’s mean at age 3 from the response scale’s theoretical midpoint was used instead of skewness.) Figure 5 displays age differences in average standard deviations for 41 lower-skewness items (absolute skewness less than 0.50 at age 3) and 53 higher-skewness items: although the trend of increasing variance is stronger for items that were initially more

skewed it is still clearly recognizable in the less skewed items, with confidence intervals at earlier ages not overlapping with those after age 7 years. To further test whether age differences in item skewness could explain age differences in variance, for each item we computed the partial (Spearman) correlation of age with age-specific item variance while controlling for age-specific item skewness. The median of these 94 partial correlations was .75 (interquartile range = .51 to .88); 72 of the correlations were above .50, while only 8 were negative (the respective numbers of items had been 82 and 5 when skewness was not controlled). Taken together, these results indicate that decreases in item skewness may have somewhat contributed to—but did not fully explain—increases in item variance at older ages.[1]

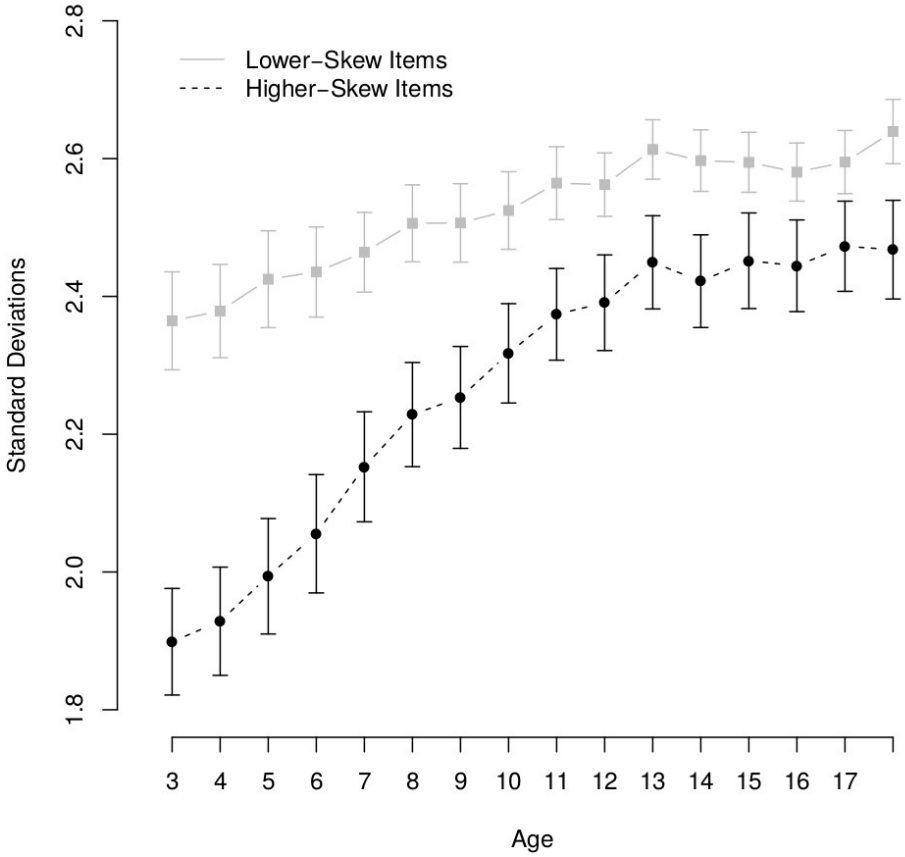

Figure 5. Average standard deviations for 41 items with absolute skewness values lower than 0.5 and 53 items with absolute skewness values 0.5 or higher (vertical lines indicate 95% confidence intervals).

Conclusion

The results of Study 1 indicate that individual differences in youth personality traits tend to increase in magnitude throughout childhood and early adolescence, and then generally plateau. Specifically, variance in all of the Little Six traits except Extraversion, variance in almost all of the specific nuance traits within the Little Six, and variance in a variety of nuance traits beyond the Little Six, all tended to increase with age. Moreover, these increases in variance could not be fully explained by methodological artifacts such as ceiling effect or mean differences, suggesting that individual differences in youth personality traits do tend to become more pronounced with age.

1 We also repeated these analyses while controlling for the items' age-specific mean responses, rather than age-specific skewness. The results tended to be similar (median correlation was .66, 57 correlations were above .50, 12 below zero), indicating that neither item skewness nor item means can fully explain age differences in item variance. Likewise, we ran the analyses separately for boys and girls: very similar pattern of increasing variance was observed in both sexes (see Online Supplementary Material).

Study 2

Study 1 found greater variance in youth personality traits at older ages. Study 2 was conducted to test whether this developmental tend generalizes to a new sample, to a new personality measure, to a new level of abstraction—facet traits, which are narrower than the Little Six domains but broader than item-level nuance traits—and to a new cultural context: the Asian part of Russia.

Method

Sample and Procedure

Participants in Study 2 were parents recruited in Novosibirsk (Russia's third largest city located in Siberia) and nearby rural areas. As described by Slobodskaya (2014), these parents were approached via child care centers and schools, or in person, including home visits. Among other measures, parents were asked to complete the Inventory of Child Individual Differences-Short version (ICID-S). The present study used only ICID-S data (valid observations for at least 10 items). The target sample consisted of 1996 children (48% female) ranging in age from 2 through 18 years old (M = 10.3, SD = 4.3). The children were grouped into five age groups: 2 to 6 years (N = 491), 7 to 9 years (N = 357), 10 to 12 years (N = 402), 13 to 15 years (N = 494), and 16 to 18 years (N = 252); these group boundaries are admittedly arbitrary, but we deemed them to be most appropriate to cover all ages with subsamples sufficiently large for stable parameter estimates. Most data came from mothers and most of the children lived with both biological parents. Researchers aimed for as large a sample as possible. The study was approved by the State Research Institute of Physiology and Basic Medicine SB RAMS Ethics Committee.

Measure

The Inventory of Child Individual Differences-Short version (ICID-S; Slobodskaya & Zupancic, 2010) is an age and culture neutral instrument designed to assess child personality. The ICID-S for parents was developed from the full instrument (Halverson et al., 2003), and maintains levels of reliability and validity comparable to the full inventory. The ICID-S version used in the present research includes 52 items measuring 15 facet-level traits: Achievement Orientation, Activity Level, Antagonism, Compliant, Considerate, Distractible, Fearful, Intelligent, Negative Affect, Openness to Experience, Organized, Positive Emotions, Shy, Sociable and Strong Willed. Each item was rated on a scale ranging from 1 (*much less than the average child*) to 9 (*much more than in the average child*). In the present study, alphas for the ICID-S scales ranged from .68 to .86 with a median of .73. For comparability with Study 1, we grouped the 15 ICID-S facets in terms of the Little Six domains on the basis of previous theory and research (Deal, Halverson, Martin, Victor, & Baker, 2007; Halverson et al., 2003; Shiner & DeYoung, 2013; Soto, 2016; Soto & John, 2014).

Data analysis

Quantitative analyses were carried out using R statistical language. The scripts and data are made publicly available in Online Supplementary Material (osf.io/aeu2b).

Results and Discussion

Preliminary Analyses

In contrast with Study 1, where many items had very skewed distributions, the skewness of the ICID-S items was generally low; absolute values ranged from 0 to 0.65 with a mean of 0.14, and only 2 items had skewness values over 0.50. For facet scores, skewness estimates ranged from 0.01 to 0.33 with median of .16. A visual inspection of item and facet score distributions in all age groups also confirmed that ceiling/floor effects were generally not a concern for ICID-S data (see Online Supplementary Material).

As the next step, we tested for measurement invariance of the 15 facets across the 5 age groups (2 to 6, 7 to 9, 10 to 12, 13 to 15, and 16 to 18 years). Similarly to Study 1, unidimensional CFA models were fitted for each facet and as many item residual correlations were allowed as necessary to achieve acceptable model fit. For most facets, no residual correlations were required, but Organized facet required three, Fearful required two, and Strong Willed, Shy and Openness to Experience facets required one residual correlation each. (As in Study 1, the main findings were similar without the residual correlations; see Figure 6SA in Online Supplementary Material.)

This was followed by fitting these models using multi-group CFA. All facets but Compliant met the criterion for weak MI (equality of factor loadings); for Compliant the decrease in CFI was also relatively small (.02). All facet showed evidence for equality of residual covariances. Six facets did not meet the criterion for the equality of residual variances, although the decreases in CFI were again mostly small (.02 for Fearful, Shy, Openness to Experience and Organized, .03 for Compliant, and .04 for Activity Level). Strictly speaking, the five scales did not measure the respective constructs with the required level of invariance, but because the deviations were relatively small, the overall level of MI was considered acceptable. For the sake of completeness, we also tested for the invariance of intercepts: six facets failed to achieve invariance, although the drops in CFI were smaller (between .02 and .06) than in Study 1 (see Online Supplementary Material for details).

Age Differences in Facet and Nuance Trait Variance

Age differences in the variance of the latent facet scales are presented in Figure 6; facets are grouped according to the Little Six domain to which they purportedly pertain to (with Extraversion and Activity combined). The kernel density plots of facet scores for all age group except the oldest are given in Online Supplementary Material.

Figure 6 shows a less clear trend for increasing variance than appeared in Study 1. The upward trajectory was evident for the facets of Conscientiousness as well as for some Agreeableness, Openness and Neuroticism facets. For a number of facets, however, there was no clear pattern of age differences in variance. For example, although Activity was among the traits displaying the strongest increases in variance with age in Study 1, this was not the case in Study 2. Comparisons of the CCQ Activity and ICID-S Active item content indicated considerable similarity between the two measures (e.g., both scales directly refer to physical activity, energy, and movement).

However, the findings were consistent with Study 1 for nuance-level trends. Figure 6 (bottom-right panel) presents age differences in variance at the level nuances, as represented by individual ICID-S items. The pattern of curvilinearly increasing variance at later ages is clearly evident and the shape of the curve is strikingly similar to the findings of Study 1: variance increases into early adolescence (here ages 13 to 15) and then plateaus or even slightly decreases[2]. For example, the 95% confidence intervals of the item standard deviation means in the two oldest age groups did not overlap with the confidence intervals from the youngest group (and those of the 10- to 12-year-olds only barely overlapped with those of the youngest age group). As suggested by the facet-level analyses, the pattern of increasing personality variance was strongest for the Conscientiousness nuances. However, the pattern was still clearly observed when Conscientiousness was omitted from the nuance-level analysis (bottom-right panel of Figure 6).

Age differences in personality variance appearing more pronounced at the nuance level than at the facet level suggests that the variance that tended to increase with age was often the unique personality variance that nuance-level traits captured, but latent trait scores filtered out.

2 It is unlikely that these trends were driven by increasing mean levels because, on average, item scores slightly declined with age (from 4.19 to 4.12).

Indeed, in the multi-group CFA with loading and residual covariances equality constraints but no residual variance constraints the residual variances tended to be higher in the older age groups as opposed to the youngest group (with a dip in the oldest group; average standard deviations across all item residual variances and the 95% confidence intervals of these averages are given in Figure 7S in the Online Supplementary Material). We also calculated age group differences in the variance of observed (not latent) ICID-S facet scores (see Figure 6SB in Online Supplementary Material) and these analyses showed somewhat clearer tendency for increasing variance in older age groups, although not for all facets. This is also consistent with the hypothesis that increases in variance often pertained to nuance-level traits, because observed scale scores (which aggregate *both* items' shared and unique variance) captured nuance-level effects better than latent trait scores (which *only* captured their shared variance).

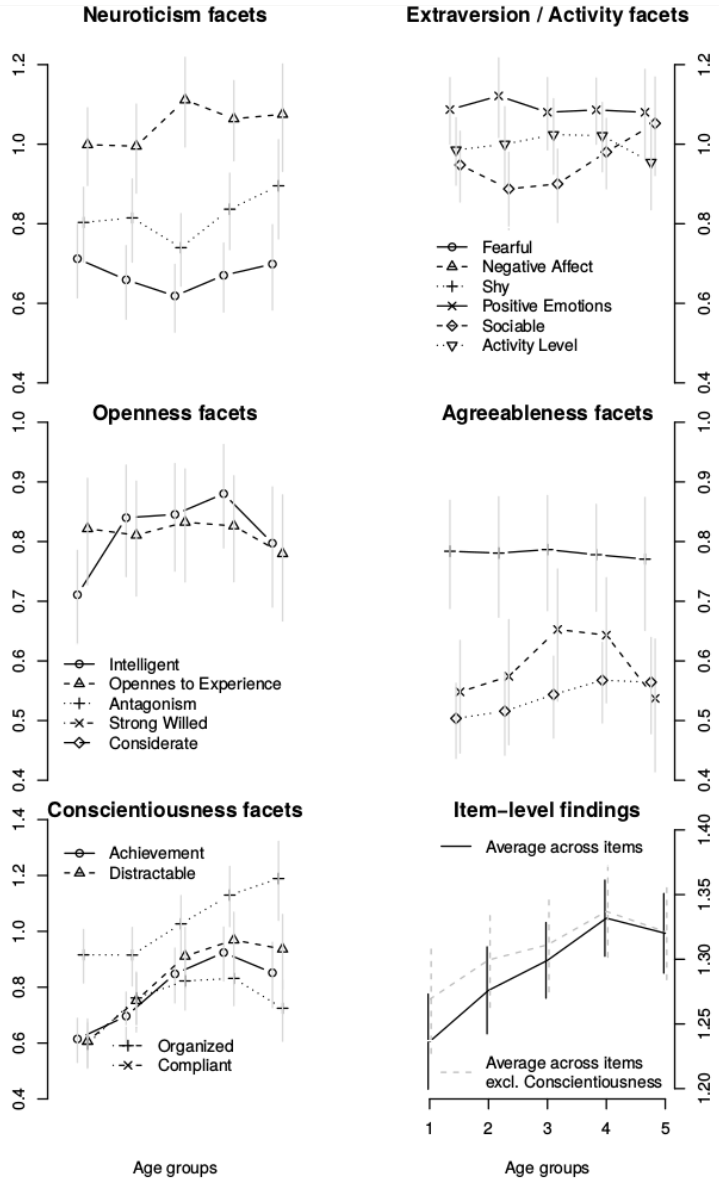

Figure 6. Standard deviations of the 15 latent facet scores of the Russian version of the ICID-S and average standard deviations of their items for five age groups (2 to 6, 7 to 9, 10 to 12, 13 to 15, and 16 to 18 years; vertical lines indicate 95% confidence intervals). The facets are grouped according to the FFM traits.

Similarly to Study 1, we additionally sought to fit a single CFA model simultaneously for all 15 latent traits, but this model could not be properly identified as either a single-group or

multiple-group model, regardless of the level of parameter equality constraints or the presence of residual correlations.

Conclusion

Study 2 was based on a relatively smaller sample than Study 1: the number of age groups was smaller and each group contained fewer children. As a result, the curves representing age differences in variance were less smooth in Study 2, especially for individual traits. The overall level of convergence across the studies, however, was notable, at least as far as nuance-level analyses are concerned. Taken together, therefore, the results of Study 2 further support the conclusion that the variance of youth personality traits tends to increase with age, especially from early childhood into early adolescence.

General Discussion

The findings of two studies, based on nearly 18,000 youths, indicate that individual differences in personality traits tend to become more pronounced from early childhood through early adolescence, after which point the magnitude of individual differences no longer increases. This pattern generalized across two independent samples, two cultural contexts and two personality measures. The pattern generalized across almost all personality characteristics in the first (English-speaking) sample, whereas it was clearly present for the most specific personality characteristics (nuances) as well as for some aggregate traits (facets) in the second (Russian) sample. This finding has both theoretical and practical implications.

Implications for Understanding Personality Development

What explains the present pattern of results, in which personality variance showed a positive age trend from early childhood through early adolescence and then a flat age trend across late adolescence (and possibly henceforth; Mõttus et al., 2016)? One possibility is that the key developmental mechanisms underlying increases in youth personality variance may operate more strongly during childhood than adolescence. For example, increases in personality variance during childhood could be primarily due to rapid expansions of youths' behavioral repertoires as new cognitive (Flavell et al., 1993; Inhelder & Piaget, 1958), emotional (Murphy et al., 1999), and self-regulatory (Demetriou, 2000; Gestsdottir & Lerner, 2008) capacities come on line through biological maturation and learning processes. By mid-adolescence, youths' available capacities may be broad enough to fully express the key aspects of their personality. Although further capacities continue to accrue through late adolescence and adulthood, they may have little or no effect on the overall clarity with which individuals can express—and thus with which parents, peers, and other observers can detect—their personality traits.

A related possibility is based on the corresponsive principle by which individuals seek out situations and social roles consistent with their personality traits, which then further reinforce those same traits (Caspi et al., 2005; Roberts et al., 2008). Specifically, it is plausible that the person-environment transactions underlying the principle operate in a non-linear fashion: the further an individual's traits are pulled from their baseline as a result of the transactions, the harder it may become to pull them yet further. This can be conceptualized as a form of gene-environment interaction: an environment that has been sought out or created to facilitate the manifestation of pre-existing genetic dispositions can only do this up to a certain level, after which the genetic dispositions become less sensitive to further environmentally driven change.

An alternative possibility is that the mechanisms underlying positive age trends in youth personality variance may continue into adolescence and adulthood, but become offset by countervailing mechanisms that pull people towards uniformity (Mõttus et al., 2016). For example, socialization pressures on behavior may intensify during adolescence, as youths are increasingly told to "grow up" and "act your age," and are generally held to higher behavioral standards in preparation for adulthood (Denissen et al., 2013). Research on mean-level age

differences in personality traits provides some support for this possibility, as levels of Agreeableness and Conscientiousness—two key targets of socialization pressure—increase substantially during late adolescence (Soto, 2016; Soto et al., 2011; van den Akker et al., 2014). However, the present finding that positive age trends in personality variance slowed or stopped during adolescence for traits subject to both relatively high (e.g., Agreeableness, Conscientiousness) and relatively low (e.g., Openness, Activity) degrees of socialization pressure may argue against this explanation (Caspi et al., 2005; Digman, 1997).

Another possible explanation for the present findings does not involve any direct environmental influences. It has been suggested that personality development may primarily reflect intrinsic maturation that happens regardless of normal environmental variability—similarly, for example, to the development of height in economically developed nations that provide adequate nutrition for most children (McCrae & Costa, 2008). In such environments, height also shows increasing variability until puberty, after which the magnitude of individual differences becomes more stable (Tanner, Whitehouse, & Takaishi, 1966). If personality traits develop similarly to height, then intrinsic maturation may represent development from a less differentiated personality profile towards a more differentiated one, which could contribute to gradual increases in individual differences—up until the age at which most children have attained their fully differentiated profiles.

If this explanation were correct, then one would expect increasing heritability estimates throughout childhood as genetic influences increasingly take over the role of shaping phenotypic variance. However, although this is consistent with observations for cognitive traits (Tucker-Drob, Briley, & Harden, 2013), the evidence is exactly the opposite for personality characteristics: their overall level of heritability tends to decline over childhood and adolescence, despite increasing stability of the remaining genetic influences (Briley & Tucker-Drob, 2014). This is not in line with an intrinsic maturation-based interpretation of the present findings. That being said, decreasing heritability may not be in keeping with the corresponsive principle-based interpretation either, because person (or gene)-environment correlations would also be expected to increase heritability estimates (Purcell, 2002). One should therefore not expect declining heritability levels as children grow older and accumulate influences resulting from person-environment transactions.

A related hypothesis is that there are individual differences in youths' rates of intrinsic personality maturation, which could also produce positive age trends in the magnitude of personality variance. At young ages, the children maturing at a faster rate cannot have been doing this for long compared to those developing more slowly, but the effects of their relatively faster developmental pace may accumulate over time, pushing them further ahead of their slower-developing peers of the same chronological age. However, if this were the primary explanation for increases in personality variance across childhood, then slower developers would be expected to catch up when development generally slows down, and personality variance would eventually decrease. Neither the present findings nor previous research indicate an eventual decrease in personality variance.

A final possibility is that the content of most personality test items may simply be more applicable to adolescents and adults than to children, and such items may therefore provide greater discrimination at older ages. This explanation may primarily apply to personality tests that have been developed to describe adults. However, both of the measures used in the present research were specifically developed to describe children. The CCQ was originally constructed from descriptions of children provided by developmental researchers and child clinicians (Block & Block, 1980), and later refined using feedback from parents of children and adolescents (Caspi et al., 1992). The ICID was developed from parents' free descriptions of children in four countries (the United States, China, Greece, and the Netherlands) and three age groups (3, 6, and 9-12 years old); its final scales were constructed from items whose content generalized across

cultures and ages (Halverson et al., 2003). Even within these youth-oriented measures, some items may be more (e.g., "Is obedient and does what they are told") or less (e.g., "Daydreams; often gets lost in thought or a fantasy world") applicable to young children. However, the finding that increases in personality variance generalized across traits and individual items suggests that this pattern was not primarily due to age-inappropriate items.

In sum, the present findings seem most consistent with the possibilities that positive age trends in personality variance across childhood are primarily due to (a) expansions of youths' behavioral capacities, through a combination of biological maturation and learning, that allow older children and adolescents to more clearly express their distinctive personality traits, and (b) person-environment transactions that reinforce youths' personality dispositions as they age—but only up to a point. Importantly, these explanations in no way contradict each other. However, we caution that this conclusion is based on the interpretation of descriptive, cross-sectional age trends, and should therefore be considered speculative. Future research is needed to distinguish between and directly test these possible developmental mechanisms.

Cross-Study Differences

The pattern of findings was clearer in Study 1, mostly based on children of English-speaking parents, than in Study 2, based on Russian children. This could be due to a number of reasons such as Study 1 having a sample eight times larger than the one employed in Study 2 or there being cross-cultural differences in personality development or how parents perceive/rate their children. The latter could well be true, but since we currently do not have a good theoretical explanation for why this might be the case we only note this possibility, without further speculations. Additionally, the differences may have resulted from different instruments or personality constructs, although the pervasiveness with which the pattern generalized across constructs in Study 1 may speak against this possibility. However, we emphasize that at the level of most specific personality traits—nuances—the pattern *did* replicate well across the two studies, with the curvilinear age-trajectory being strikingly similar.

Advancing Our Understanding of Mean-Level Trends

Studying developmental trends in the shape of personality trait distributions may provide a more refined understanding of often-described mean-level trends. For example, in the present data one of the clearest mean-level trends was for Activity, which, on average, declined throughout childhood and adolescence (Slobodskaya & Akhmetova, 2010; Soto, 2016). In Study 1, based on a very large sample of children, an examination of age differences in the distribution of Activity scores (Figures 2 and 3) revealed that the trend was at least partly (until early puberty) due to increasing variance—in particular, due to an increasing prevalence of lower values. This could suggest that in this predominantly North-American sample decline rates are stronger among children with relatively lower initial levels of the trait, whereas those with relatively higher Activity scores could continue to be among the most active individuals; perhaps they are less sensitive to whatever pressures generally tend to reduce the level of this trait. In contrast, Extraversion (Study 1) and Activity and Positive Emotions (Study 2) showed a pattern of mean-level decline (Slobodskaya & Akhmetova, 2010; Soto, 2016), but did not show greater variance at older ages (Figures 1 and 6). These contrasting patterns may suggest that the similar mean-level trends for Activity versus Extraversion in Study 1, and Activity and Positive Emotions in Study 2 could be attributable to different developmental mechanisms.

Moreover, even if there is little in the way of systematic mean-level trends but systematic increases in variance, such as was the case for Conscientiousness (Figures 1 and 6; Slobodskaya & Akhmetova, 2010; Soto, 2016), this may suggest that the interpretation of the former might need to be qualified by the latter. Specifically, small or non-existent differences in mean scores do not necessarily indicate that people, on average, do not show age differences in the trait: in fact, most people *might* display age differences but their direction depends on where the

particular person is in the distribution of the trait. High-scorers may tend in the direction of scoring yet higher, whereas the opposite might be true for those at the lower end of the distribution. We note, however, that due to poor levels of strong MI in these data, and in such data more generally (e.g., Mõttus et al., 2015), mean-level differences in aggregate trait scores should be interpreted cautiously.

Of course, testing such possibilities would require longitudinal data—the present findings only suggest these hypotheses. So far, longitudinal investigations into mean-level changes have not paid much attention to the changes in the shapes of the distributions of personality scores. Until such investigations are carried out, mean-level differences across groups, or changes within groups over time, should not be automatically interpreted as describing the development of the "average" individual—they may pertain to changes in particular subsets of individuals, or in some instances even to no individuals at all. This is consistent with studies based on mixture modeling that have identified subgroups of individuals developing along different trajectories (Klimstra, Hale, Raaijmakers, Branje, & Meeus, 2010; Johnson, Hicks, McGue, & Iacono, 2007).

Beyond personality development

Age differences in variance may thus be informative with respect to the mechanisms and trajectories of personality development. However, the present findings are also valuable beyond these implications. First, to the extent that personality is defined as an individual's distinctive, characteristic pattern of cognition, emotion, and behavior (e.g., Funder, 2016), understanding the very extent of individual differences in these patterns should be a fundamental goal of personality psychology. Personality has been traditionally conceptualized as an adult phenomenon, but a growing body of research has applied and extended personality concepts into childhood and adolescence (Soto & Tackett, 2015). The present findings provide some support for both the traditional, adult-focused view of personality and the contemporary, youth-inclusive view: there are individual differences in youths' patterns of behavior, and these individual differences become even more pronounced by adulthood. Importantly, however, the present findings would suggest that at least some aspects of personality development happen quite early in childhood—perhaps earlier than most studies into personality development consider. Specifically, whatever drives individual differences in personality to their adult-like magnitude operates mostly in before adolescence.

Second, from a methodological point of view, all attempts to identify the correlates and consequences of personality characteristics depend on the amount of observable personality variance. Therefore, the present findings suggest that associations of personality traits with other variables may be more constrained by restricted personality variance during childhood than during adolescence and adulthood. This implication highlights the value of large-sample studies of youth personality development. Although psychological researchers are increasingly recognizing the importance of sample size and statistical power for generating replicable scientific findings (Open Science Collaboration, 2015), the pattern found here of less personality variance at younger ages suggests that this issue may be particularly crucial for detecting youth personality effects that may be somewhat attenuated by restricted variance.

Strengths, Limitations, and Future Directions

Two key strengths of the present research were (a) its use of two large samples from different cultures tested with different youth personality measures, and (b) its use of personality parent-reports. The former strength enhances the precision and generalizability of the present findings, whereas the latter helps distinguish age differences in personality variance from measurement artifacts such as age differences in reporting bias. Although parents of 18-year-olds are generally older than parents of three-year-olds, both sets of parents should be about equally capable of providing reliable and valid personality reports. On the other hand, as

children become older and spend more time outside the home, parents' ability to observe the full range of their behavior may become increasingly limited. Moreover, it may be exactly outside the home that children continue to broaden their behavioral repertoire. Thus, our reliance on parent-reports may have produced an underestimation of personality variance in older children and adolescents. This may at least partly explain the finding that variance increases generally stopped by mid-adolescence, among the reasons discussed above. Furthermore, it is possible that parents' ratings of children at different ages may be differentially biased (e.g., there are stronger personality stereotypes regarding very young children and/or teen-agers than for other stages of childhood). This is consistent with the less-than-ideal measurement invariance across age groups that we observed. Future studies examining personality development could benefit from obtaining youth personality reports from multiple sources, including teachers, multiple parents, and researchers themselves.

A second limitation is that distributions of the scale scores were sometimes truncated in Study 1, which could have confounded the results. Indeed, increases in variance appeared smaller in the reduced versions of the scales that included less skewed items—but they were still observable. Also, items that were initially more skewed tended to show somewhat larger increases in variance with age, but skewness did not fully explain age differences in item variance. Moreover, item skewness, and therefore truncated distributions, were much less evident in Study 2, and this study still replicated the key finding that personality variance tended to increase with age—most clearly for nuances.

Finally, the present research was cross-sectional, whereas longitudinal studies can further enhance our understanding of youth personality variance. Apart from allowing researchers to rule out differential sampling or cohort effects, longitudinal designs allow them to identify which particular individuals tend to move in which particular directions. For example, the present findings suggest the possibility that highly active children remain active whereas relatively low-activity children are more likely to become even less active—at least in English-speaking populations. This hypothesis and others like it can be directly tested using longitudinal data. Indeed, several studies have identified subgroups of children developing along different trajectories (Klimstra, Hale, Raaijmakers, Branje, & Meeus, 2010; Johnson, Hicks, McGue, & Iacono, 2007), although these studies have not necessarily addressed changes in the overall magnitude of individual differences.

Conclusion

The variability of youth personality traits tends to increase with age. Put differently, as children grow they tend to become less alike in terms of their personality characteristics. This key finding advances our understanding of how and why youth personality development occurs. First, mean-level age differences (or lack of them) may sometimes be driven by a relative increase in the number of individuals falling at either high or low levels of the trait (or on both ends). Second, the overall pattern of increasing trait variance is consistent with several possible developmental mechanisms, including youths' expanding behavioral capacities and person-environment transactions that reinforce (up to a point) dispositional traits. However, we note that the current study was not intended as an attempt to validate any particular mechanism: it set out to explore age trajectories in the magnitude of individual difference and discuss these in the light of relevant theoretical propositions. In conclusion, we suggest that investigations of variance become as integral a part of studying personality development as are investigations of mean-level change/differences and the stability of individual differences.

References

Aa, N. van der, Overbeek, G., Engels, R. C. M. E., Scholte, R. H. J., Meerkerk, G.-J., & Eijnden, R. J. J. M. V. den. (2009). Daily and compulsive internet use and well-being in adolescence: A diathesis-stress model based on Big Five personality traits. *Journal of Youth and Adolescence*, *38*, 765. https://doi.org/10.1007/s10964-008-9298-3

van den Akker, A. L., Deković, M., Asscher, J., & Prinzie, P. (2014). Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity principle and bidirectional associations with parenting. *Journal of Personality and Social Psychology, 107,* 736–750. https://doi.org/10.1037/a0037248

Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality*, *18*, 445–462. https://doi.org/10.1002/per.524

Block., J., & Block, J. H. (1980). *The California Child Q-set.* Palo Alto, CA: Consulting Psychologists Press.

Briley, D. A., & Tucker-Drob, E. M. (2014). Genetic and environmental continuity in personality development: A meta-analysis. *Psychological Bulletin*, *140*, 1303–1331. http://doi.org/10.1037/a0037091

Byrne, B., & Shavelson, R. (1996). On the structure of social self-concept for pre-, early, and late adolescents: A test of the Shavelson, Hubner, & Stanton (1976) model. *Journal of Personality and Social Psychology, 70,* 599-613.

Caspi, A., Block, J., Block, J. H., Klopp, B., Lynam, D., Moffitt, T. E., & Stouthamer-Loeber, M. (1992). A "common-language" version of the California Child Q-Set for personality assessment. *Psychological Assessment*, *4*, 512–523. http://doi.org/10.1037/1040-3590.4.4.512

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453–484. http://doi.org/10.1146/annurev.psych.55.090902.141913

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-git indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255. http://doi.org/10.1207/S15328007SEM0902_5

Clark, L., & Watson, D. (2008). Temperament: An organizing paradigm for trait psychology. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 265-286). New York, NY: Guilford.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development, 48,* 1-124.

Demetriou, A. (2000). *Organization and development of self-understanding and self-regulation: Toward a general theory.* In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), Handbook of self-regulation (pp. 209 -251). San Diego, CA: Academic Press.

De Pauw, S. S. W., & Mervielde, I. (2010). Temperament, personality and developmental psychopathology: A review based on the conceptual dimensions underlying childhood traits. *Child Psychiatry and Human Development*, *41*, 313–329. http://doi.org/10.1007/s10578-009-0171-8

De Pauw, S. S. W., Mervielde, I., & Van Leeuwen, K. G. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *Journal of Abnormal Child Psychology*, *37*, 309–325. http://doi.org/10.1007/s10802-008-9290-0

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*, 1246–1256.

Donahue, E. M. (1994). Do children use the Big Five, too? Content and structural form in personality description. *Journal of Personality, 62,* 45-66.

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the big five across the life span: Evidence from two national samples. *Psychology and Aging*, *23*, 558–566. http://doi.org/10.1037/a0012897

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*, 155–174.

Eisenberg, N., Carlo, G., Murphy, B., & Van Court, P. (1995). Prosocial development in late adolescence: A longitudinal study. *Child Development*, *66*, 1179–1197.

Eisenberg, N., & Morris, A. S. (2004). Moral cognitions and prosocial responding in adolescence. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 155-188). Hoboken, NJ: Wiley.

Erikson, E. H. (1968). Identity: Youth and crisis. New York, NY: Norton.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development.* London: Pearson.

Funder, D. C. (2016). *The personality puzzle* (7th ed.). New York: Norton.

Gestsdottir, S., & Lerner, R. M. (2008). Positive development in adolescence: The development and role of intentional self-regulation. *Human Development*, *51*, 202–224. https://doi.org/10.1159/000135757

Greenberger, E. (1982). Education and the acquisition of psychosocial maturity. In D. C. McClelland (Ed.), *The development of social maturity* (pp. 155-189). New York, NY: Irvington.

Greenberger, E. (1984). Defining psychosocial maturity in adolescence. In P. Karoly & J. Steffen (Eds.), *Adolescent behavior disorders: Foundations and contemporary concerns* (pp. 54- 81). Lexington, MA: Heath.

de Haan, A. D., Deković, M., van den Akker, A. L., Stoltz, S. E. M. J., & Prinzie, P. (2013). Developmental personality types from childhood to adolescence: Associations with parenting and adjustment. *Child Development*, *84*, 2015–2030. https://doi.org/10.1111/cdev.12092

Halverson, C. F., Havill, V. L., Deal, J., Baker, S. R., Victor, J. B., Pavlopoulos, V., … Wen, L. (2003). Personality atructure as serived from parental ratings of free sescriptions of children: The Inventory of Child Individual Differences. *Journal of Personality*, *71*, 995–1026. http://doi.org/10.1111/1467-6494.7106005

Harter, S. (1999). *The construction of the self: A developmental perspective.* New York, NY: Guilford.

Harter, S. (2006). The self. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 505-570). Hoboken, NJ: Wiley.

Harter, S., & Monsour, A. (1992). Developmental analysis of conflict caused by opposing attributes in the adolescent self-portrait. *Developmental Psychology, 28,* 251-260.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Johnson, W., Hicks, B. M., McGue, M., & Iacono, W. G. (2007). Most of the girls are alright, but some aren't: Personality trajectory groups from ages 14 to 24 and some associations with outcomes. *Journal of Personality and Social Psychology*, *93*, 266–284. https://doi.org/10.1037/0022-3514.93.2.266

Klimstra, T. A., Hale, W. W., Raaijmakers, Q. A. W., Branje, S. J. T., & Meeus, W. H. J. (2010). A developmental typology of adolescent personality. *European Journal of Personality, 24*, 309–323. https://doi.org/10.1002/per.744

Krueger, R. F., & Johnson, W. (2008). Behavioral genetics and personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 287-310). New York, NY: Guilford.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81,* 417-430.

Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology, 95,* 687-706. https://doi.org/10.1037/0022-0663.95.4.687

McAdams, D. P., & McLean, K. C. (2013). Narrative Identity. *Current Directions in Psychological Science*, *22*, 233–238. https://doi.org/10.1177/0963721413475622

McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review*, *19*, 97–112. http://doi.org/10.1177/1088868314541857

McCrae, R. R., & Costa, P. T. (2008). The Five-Factor Theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research (3rd ed.).* (pp. 159–181). New York, NY US: Guilford Press.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175–215. http://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543. http://doi.org/10.1007/BF02294825

Montemayor, R., & Eisen, M. (1977). The development of self-conceptions from childhood to adolescence. Developmental Psychology, 13, 314-319.

Mõttus, R., Allik, J., Hřebíčková, M., Kööts-Ausmees, L., & Realo, A. (2016). Age differences in the variance of personality characteristics. *European Journal of Personality*, *30*, 4–11. http://doi.org/10.1002/per.2036

Mõttus, R., Kandler, C., Bleidorn, W., Riemann, R., & McCrae, R. R. (2017). Personality traits below facets: The consensual validity, longitudinal stability, heritability, and utility of personality nuances. *Journal of Personality and Social Psychology, 112,* 474-490. http://psycnet.apa.org/doi/10.1037/pspp0000100

Mõttus, R., Realo, A., Allik, J., Esko, T., Metspalu, A., & Johnson, W. (2015). Within-trait heterogeneity in age group differences in personality domains and facets: Implications for the development and coherence of personality traits. *PLoS ONE*, *10*, e0119667. http://doi.org/10.1371/journal.pone.0119667

Murphy, B. C., Eisenberg, N, Fabes, N. A., Shepard, S., & Guthrie, I. K. (1999). Consistency and change in children's emotionality and regulation: A longitudinal study. *Merrill-Palmer Quarterly*, *45*, 413-444.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological. *Science*, *349*, aac4716. http://doi.org/10.1126/science.aac4716

Patterson, G. R., & Stouthamer-Loeber, M. (1984). The correlation of family management practices and delinquency. *Child Development, 55,* 1299-1307.

Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research*, *5*, 554–571. http://doi.org/10.1375/136905202762342026

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*, 1–25. http://doi.org/10.1037/0033-2909.132.1.1

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: origins and outcomes. *Journal of Personality and Social Psychology*, *78*, 122–135.

Shiner, R., & DeYoung, C. G. (2013). The structure of temperament and personality traits: A developmental perspective. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology: Vol. 2. Self and other* (pp. 113–141). New York: Oxford University Press.

Slobodskaya, H. R. (2014). The hierarchical structure of personality and common psychopathology in childhood. *Journal of Research in Personality*, *53*, 36–46. http://doi.org/10.1016/j.jrp.2014.08.005

Slobodskaya, H. R., & Akhmetova, O. A. (2010). Personality development and problem behavior in Russian children and adolescents. *International Journal of Behavioral Development, 34,* 441-451.

Slobodskaya, H. R., & Zupancic, M. (2010). Development and validation of the Inventory of Child Individual Differences-Short Version in two Slavic countries. *Studia Psychologica*, *52*, 23.

Small, B. J., Hertzog, C., Hultsch, D. F., & Dixon, R. A. (2003). Stability and change in adult personality over 6 years: findings from the Victoria Longitudinal Study. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *58*(3), P166-176.

Soto, C. J. (2016). The Little Six personality dimensions from early childhood to early adulthood: Mean-level age and gender differences in parents' reports. *Journal of Personality, 84,* 409-422. http://doi.org/10.1111/jopy.12168

Soto, C. J., & John, O. P. (2014). Traits in transition: the structure of parent-reported personality traits from early childhood to early adulthood. *Journal of Personality*, *82*, 182–199. http://doi.org/10.1111/jopy.12044

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718–737. https://doi.org/10.1037/0022-3514.94.4.718

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*, 330–348. http://doi.org/10.1037/a0021717

Soto, C. J., & Tackett, J. L. (2015). Personality traits in childhood and adolescence structure, development, and outcomes. *Current Directions in Psychological Science*, *24*, 358–362. http://doi.org/10.1177/0963721415589345

Steinberg, L., & Silverberg, S. B. (1986). The vicissitudes of autonomy in early adolescence. *Child Development, 57,* 841-851.

Tanner, J. M., Whitehouse, R. H., & Takaishi, M. (1966). Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. I. *Archives of Disease in Childhood*, *41*, 454–471. http://doi.org/10.1136/adc.41.219.454

Tucker-Drob, E. M., Briley, D. A., & Harden, K. P. (2013). Genetic and environmental influences on cognition across development and context. *Current Directions in Psychological Science*, *22*, 349–355. http://doi.org/10.1177/0963721413485087

van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*, 486–492. http://doi.org/10.1080/17405629.2012.686740